

Canonical forms of Self-Adjoint Operators in Minkowski Space-time

Krishan Rajaratnam*

April 8, 2014

Abstract

We classify the possible Jordan canonical forms of self-adjoint operators in Minkowski space-time (in fact in pseudo-Euclidean space, i.e. an indefinite inner product space) and we show how to obtain a Jordan canonical basis which also puts the metric in a canonical form.

Contents

List of Results	ii
1 Introduction	1
2 Preliminary Results	2
2.1 Pseudo-Euclidean Spaces	2
2.2 Finite Dimensional Operator Theory	4
3 Existence of the metric-Jordan canonical form	6
4 Uniqueness of the metric-Jordan canonical form	12
5 Minkowski Space	16
References	18

*Department of Applied Mathematics, University of Waterloo, Canada; e-mail: k2rajara@uwaterloo.ca

List of Results

Theorem 2.3	5
Theorem 2.5	Sylvester's Law of Inertia	6
Theorem 3.1	Fundamental Properties of Self-Adjoint Operators	7
Theorem 3.7	Existence of the metric-Jordan canonical form [O'N83]	11
Theorem 4.2	Uniqueness of the metric-Jordan canonical form	13
Theorem 4.5	Isometric Equivalence of self-adjoint operators	15
Theorem 5.1	Properties of self-adjoint operators in Minkowski space	17

1 Introduction

Self-adjoint operators are ubiquitous in pseudo-Riemannian geometry and hence in relativity theory as well. Any linear operator metrically equivalent to a symmetric contravariant tensor is self-adjoint. The Ricci tensor, the Hessian of a smooth function, the shape operator associated with a pseudo-Riemannian hypersurface or an umbilic pseudo-Riemannian submanifold [O’N83, Definition 4.18], and Killing and conformal Killing tensors are all examples of such tensors. In relativity theory the energy-momentum tensor is one as well.

Most of our notions come from the theory of inner products, the difference being here that we weaken the positive definite condition on the inner product to just indefinite¹ (i.e. non-degenerate). We refer to such a product as a scalar product [O’N83], and a vector space equipped with a scalar product a scalar product space. The definitions of self-adjoint, unitary, normal etc. hold in this more general context as well. So we say a linear operator T on vector space V with scalar product $\langle \cdot, \cdot \rangle$ is self-adjoint if $\langle Tx, y \rangle = \langle x, Ty \rangle$ for all $x, y \in V$. The problem we solve here is that of finding a simultaneous canonical form for a self-adjoint operator T and the scalar product $\langle \cdot, \cdot \rangle$.

This problem has some equivalent formulations that are worth noting. It often appears in the literature as finding canonical forms for “pairs of (real) symmetric matrices (bilinear forms)” under congruence where one is assumed to be non-singular (i.e. invertible). This equivalence can be observed if we fix a scalar product $\langle \cdot, \cdot \rangle$ and a self-adjoint operator T , then define another bilinear form $[\cdot, \cdot]$ by $[x, y] = \langle Tx, y \rangle$. Then the problem of finding a simultaneous canonical form for T and $\langle \cdot, \cdot \rangle$ is equivalent to that of finding one for the symmetric bilinear forms $\langle \cdot, \cdot \rangle$ and $[\cdot, \cdot]$ (or their matrix representations under congruence). Also a symmetric bilinear form is equivalent to a quadratic form by the polarization identity; this gives another formulation of the problem.

The problem of obtaining the possible canonical forms for self-adjoint operators in Minkowski space-time and more generally in a scalar product space has been addressed many times. A solution appeared in [Bro06] where the problem was formulated as that of finding canonical forms for pairs of quadratic forms. Also see [GL05, Chapter 5] for another more contemporary solution, historical notes and references. The solution in [GL05] is for a more general setting where the scalar product is a sesquilinear form. It also contains a classification of unitary operators and much more related results. Also see [LR05] and references therein; in [LR05] the Kronecker canonical form is used to solve a more general problem.

Even though the results are out there, they are relatively hard to find and often outdated. Our approach to this problem is based on O’Neil’s solution given in exercises 18-19 in [O’N83, P. 260-261]. Motivated by the Jordan canonical form, we develop an algorithm to find a Jordan canonical basis for a self-adjoint operator which also gives a canonical form for the scalar product. Our derivation has some advantages: it only depends on some results from Jordan form theory, we are able to prove existence

¹We will define these notions more precisely in the next section.

and uniqueness of the canonical form independently of the corresponding results from Jordan form theory (i.e. ours results have less dependencies), and we obtain a simple algorithm to calculate the canonical forms for self-adjoint operators. The draw back is that our solution is less general than others.

2 Preliminary Results

2.1 Pseudo-Euclidean Spaces

Suppose V is a vector space over a field \mathbb{F} (which for us is \mathbb{R} or \mathbb{C}). A symmetric bilinear form on V is bilinear function $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{F}$ such that $\langle x, y \rangle = \langle y, x \rangle$ for all $x, y \in V$. A symmetric bilinear form $\langle \cdot, \cdot \rangle$ is called non-degenerate if for a fixed $x \in V$, $\langle x, y \rangle = 0$ for all $y \in V$ implies $x = 0$.

Given a non-zero vector $x \in V$, it is classified as follows:

timelike If $\langle x, x \rangle < 0$

lightlike If $\langle x, x \rangle = 0$

spacelike If $\langle x, x \rangle > 0$

We define a scalar product on a vector space V as a non-degenerate symmetric bilinear form on V . A vector space V equipped with a scalar product is called a scalar product space. The index of a real scalar product space V , denoted $\text{ind } V$, is defined as the number of timelike basis vectors in an orthogonal basis for the scalar product, which is an invariant of the scalar product by Sylvester's law of inertia. For all notions related to the index, we will assume the scalar product space is real.

The Euclidean metric given as follows is an example of a non-degenerate scalar product:

$$\langle x, y \rangle = \sum_{i=1}^n x_i y_i$$

A vector space equipped with the Euclidean metric is called a Euclidean space. The standard example of a non-degenerate scalar product with non-zero index is the Minkowski metric given as follows:

$$\langle x, y \rangle = \sum_{i=2}^n x_i y_i - x_1 y_1$$

A vector space equipped with the Minkowski metric is called a Lorentz space (Minkowski space).

Given a subspace $H \subseteq V$, we denote the orthogonal subspace of H as H^\perp which is defined as follows:

$$H^\perp = \{x \in V : \langle x, y \rangle = 0 \text{ for all } y \in H\}$$

H^\perp is complementary to H (i.e. $V = H \oplus H^\perp$) iff the restriction of the scalar product to H is non-degenerate [O'N83, P. 49]. One can also show that for a non-degenerate subspace H , $\text{ind } V = \text{ind } H + \text{ind } H^\perp$. If U, W are subspaces of V , then $V = U \oplus W$ means that $V = U \oplus W$ and $U \perp W$.

A linear operator T on a scalar product space V is said to be self-adjoint if $\langle Tx, y \rangle = \langle x, Ty \rangle$ for all $x, y \in V$. Note that any polynomial in T is self-adjoint if T is.

We denote the complexification of a real vector space V by $V^\mathbb{C}$. We define the complexified bilinear form to be the symmetric bilinear form in $V^\mathbb{C}$ obtained from the real one by a linear extension. Note that the complexified bilinear form is symmetric, in contrast with the usual Hermitian form which is not symmetric. It follows immediately from the definition that the complexified bilinear form is non-degenerate iff the real bilinear form is non-degenerate. Thus a real scalar product space, V , can be canonically complexified to a complex scalar product space, hereafter denoted $V^\mathbb{C}$.

A set v_1, \dots, v_n for V is said to be *orthonormal* if $\langle v_i, v_i \rangle = \pm 1$ and $\langle v_i, v_j \rangle = 0$ for $i \neq j$. Clearly an orthonormal set forms a basis for V and the metric in this basis is $g = \text{diag}(\pm 1, \dots, \pm 1)$.

Now assume the scalar product $\langle \cdot, \cdot \rangle$ is (possibly) degenerate. We say a sequence of vectors v_1, \dots, v_p is a *skew-normal sequence* of (length p) and (sign $\varepsilon = \pm 1$) if $\langle v_i, v_j \rangle = \varepsilon$ when $i + j = p + 1$ and $\langle v_i, v_j \rangle = 0$ otherwise. We will show shortly that these vectors are necessarily linearly independent, so let $H = \text{span}\{v_1, \dots, v_p\}$. Then the bilinear form restricted to H is skew-diagonal and is given as follows:

$$g = \begin{pmatrix} 0 & & & \varepsilon \\ & \ddots & & \\ & & \ddots & \\ \varepsilon & & & 0 \end{pmatrix}$$

The following lemma shows that a skew-normal sequence forms a linearly independent set and it gives the index of the space spanned by such a set of vectors.

Lemma 2.1

Suppose V is a (possibly complex) scalar product space. Suppose $\varepsilon \in \{-1, 1\}$ and let z_1, \dots, z_p be a skew-normal sequence of sign ε .

Then z_1, \dots, z_p form a linearly independent set and the subspace H spanned by these vectors is non-degenerate and has index:

$$\text{ind } H = \begin{cases} \lfloor \frac{p+1}{2} \rfloor & \text{if } \varepsilon = -1 \\ p - \lfloor \frac{p+1}{2} \rfloor & \text{if } \varepsilon = 1 \end{cases} \quad \square$$

PROOF Given $1 \leq i \leq p$, denote the additive conjugate of i by $i' = p + 1 - i$. Note that $1 \leq i' \leq p$ and $i + i' = p + 1$. Suppose $i < \frac{p+1}{2}$ and $j > \frac{p+1}{2}$. Define vectors v_i and

v_j as follows:

$$\begin{aligned} v_i &= \frac{1}{\sqrt{2}}(z_i + z_{i'}) \\ v_j &= \frac{1}{\sqrt{2}}(z_j - z_{j'}) \end{aligned}$$

If $2i = p + 1$ then let $v_i = z_i$. Now for $i < j$ suppose $i + j = p + 1$, then note that $i' = j$ and $j' = i$, observe that:

$$\begin{aligned} \langle v_i, v_j \rangle &= \frac{1}{2}(\langle z_i + z_j, z_j - z_i \rangle) \\ &= \frac{1}{2}(\langle z_j, z_j \rangle + \langle z_i, z_i \rangle) \\ &= 0 \end{aligned}$$

$$\begin{aligned} \langle v_i, v_i \rangle &= \frac{1}{2}(\langle z_j + z_i, z_j + z_i \rangle) \\ &= \frac{1}{2}(2 \langle z_i, z_j \rangle) \\ &= \epsilon \end{aligned}$$

$$\begin{aligned} \langle v_j, v_j \rangle &= \frac{1}{2}(\langle z_j - z_i, z_j - z_i \rangle) \\ &= -\frac{1}{2}(2 \langle z_i, z_j \rangle) \\ &= -\epsilon \end{aligned}$$

Furthermore if $2i = p + 1$ then $\langle v_i, v_i \rangle = \epsilon$. Now suppose $i + j \neq p + 1$, then the only way in which $\langle v_i, v_j \rangle \neq 0$ if $i + j' = p + 1$ or $i' + j = p + 1$, but one can see immediately that $i + j' = p + 1$ iff $i' + j = p + 1$. So suppose $i + j' = p + 1$. Then $p + 1 = i + i' = i + p + 1 - j$, hence $i = j$, in which case $\langle v_i, v_j \rangle$ reduces to the ones examined.

Thus we conclude that $\langle v_i, v_i \rangle = \epsilon$ if $i \leq \lfloor \frac{p+1}{2} \rfloor$, $\langle v_i, v_i \rangle = -\epsilon$ if $i > \lfloor \frac{p+1}{2} \rfloor$ and $\langle v_i, v_j \rangle = 0$ if $i \neq j$. Thus the conclusions follow. \blacksquare

2.2 Finite Dimensional Operator Theory

Given a complex scalar λ , a non-zero vector $x \in V$ is called a generalized eigenvector for T corresponding to λ if $(T - \lambda I)^p x = 0$ for some positive integer p .

Definition 2.2 (Generalized Eigenspaces)

Let T be a linear operator on finite-dimensional complex vector space V and let λ be an eigenvalue of T . The generalized eigenspace (g-space) corresponding to λ , denoted K_λ , is the subset of V defined by:

$$K_\lambda = \{x \in V : (T - \lambda I)^p(x) = 0 \text{ for some positive integer } p\} \quad \square$$

We say a set of distinct scalars $\lambda_1, \dots, \lambda_k$ are the spectrum of T if they constitute all eigenvalues of T . Furthermore the kernel of an operator T is denoted by $\ker T$ or $N(T)$.

The following results concerning the g-spaces of a linear operator are proven in [FIS03, Section 7.1].

Theorem 2.3

Let T be a linear operator on finite-dimensional complex vector space V . Suppose λ and μ are distinct eigenvalues of T , then the following statements are true:

1. K_λ is a non-zero T invariant subspace of V
2. $K_\lambda \cap K_\mu = \{0\}$
3. Let $U = (T - \lambda I)|_{K_\mu}$, then K_μ is $(T - \lambda I)$ -invariant and U is a bijection.
4. If m is the algebraic multiplicity of λ then $K_\lambda = N(T - \lambda I)^m$ and $\dim K_\lambda \leq m$.
5. If $\lambda_1, \dots, \lambda_k$ is the spectrum of T , then $V = \bigoplus_{i=1}^k K_{\lambda_i}$ \square

Hence the above theorem implies T is block diagonal in a basis adapted to the g-spaces.

Definition 2.4

Let T be a linear operator on finite-dimensional complex vector space E and let x be a generalized eigenvector of T corresponding to the eigenvalue λ . Suppose that p is the smallest positive integer such that $(T - \lambda I)^p(x) = 0$. Then the ordered set

$$\{(T - \lambda I)^{p-1}(x), (T - \lambda I)^{p-2}(x), \dots, (T - \lambda I)(x), x\}$$

is called a cycle (p -cycle) of generalized eigenvectors for T with eigenvalue λ . $(T - \lambda I)^{p-1}(x)$ and x are called the initial vector and the end vector of the cycle, respectively. We say that the cycle has length p and x generates a cycle (p -cycle) of generalized eigenvectors. \square

We first note that the subspace spanned by a p -cycle has dimension p . We also observe that a given cycle of generalized eigenvectors generated by x with eigenvalue

λ lie in K_λ . Also T restricted to this cycle has the following matrix representation:

$$\begin{pmatrix} \lambda & 1 & & & \\ & \lambda & \ddots & & 0 \\ & & \ddots & 1 & \\ & & & \lambda & 1 \\ 0 & & & & \lambda \end{pmatrix}$$

We denote $U_\lambda = T - \lambda I$ and if λ is fixed we remove the subscript and refer to $U_\lambda = U$.

Suppose T is a real linear operator and let λ be an eigenvalue with non-zero imaginary part. Suppose x generates a cycle of generalized eigenvectors of length p with eigenvalue λ whose end vector has linearly independent real and imaginary parts. Then it follows that \bar{x} generates a cycle of generalized eigenvectors of length p with eigenvalue $\bar{\lambda}$ which is linearly independent of the cycle generated by x . We denote the real subspace generated by these vectors as $K_{\lambda \oplus \bar{\lambda}}$ and call this the *real subspace* spanned by the cycle generated by x . If $\lambda \in \mathbb{R}$, then this real subspace is just K_λ .

Knowledge of the Jordan canonical form is unnecessary for our derivation. Although for readers familiar with it, note that $K_\lambda \simeq N_1 \oplus \frac{N_2}{N_1} \oplus \dots \oplus \frac{N_p}{N_{p-1}}$ where $N_i = \ker U_\lambda^i$. This shows the non-uniqueness of a given Jordan canonical basis. We will use this fact to find a Jordan canonical basis for a self-adjoint operator adapted to the scalar product.

In order to prove the uniqueness of the metric-Jordan canonical form of a self-adjoint operator we will need some theory on symmetric bilinear forms. First a *diagonal representation* of a symmetric bilinear form is a basis in which the matrix representation of the form is diagonal.

Theorem 2.5 (Sylvester's Law of Inertia)

For any symmetric bilinear form defined over a real vector space, the number of positive diagonal entries and negative diagonal entries in a diagonal representation is independent of the diagonal representation.

For any symmetric bilinear form defined over a complex vector space, the number of non-zero diagonal entries in a diagonal representation is independent of the diagonal representation. □

PROOF For the real case, see Theorem 6.38 in [FIS03] or Theorem 6.8 in [Jac12]. For the complex case, see Theorem 6.6 in [Jac12] ■

3 Existence of the metric-Jordan canonical form

In this section we will show how to obtain the canonical form, culminating in Theorem 3.7. First we need some properties of self-adjoint operators.

Theorem 3.1 (Fundamental Properties of Self-Adjoint Operators)

Suppose V is a scalar product space and T is a self-adjoint operator on V . Suppose $H \subseteq V$ is an invariant subspace of T . Then

1. $T(H^\perp) \subseteq H^\perp$, i.e. H^\perp is an invariant subspace of T .
2. $(\ker T)^\perp = \text{range } T$ and $V = \ker T \oplus \text{range } T$ iff either $\ker T$ or $\text{range } T$ is a non-degenerate subspace
3. Any polynomial in T is self-adjoint. □

PROOF The proofs are immediate. ■

Remark 3.2

The first statement of the above theorem also holds for unitary operators on V , as noted by O'Neil in [O'N83, Section 9.4]. □

The idea behind obtaining the canonical forms is as follows. First suppose that T is a self-adjoint operator on a scalar product space. When E is a Euclidean space, one can easily diagonalize T using property 1 and the fact that self-adjoint operators in Euclidean space have real eigenvalues. Indeed, after one finds a single eigenvector v , one can use property 1 to deduce that the subspace orthogonal to v must be T -invariant. Since in Euclidean space the subspace orthogonal to v must be complementary to v , one can repeat this procedure to find a basis of eigenvectors for T .

For general indefinite scalar products, our goal will be to find a cycle of generalized eigenvectors for T such that they span a non-degenerate subspace. Then as in the Euclidean case, we can use property 1 to inductively build a Jordan canonical basis for T . We will now develop a series of lemmas to show that any self-adjoint operator admits a cycle of generalized eigenvectors whose span is a non-degenerate subspace. Then we will combine these lemmas in Theorem 3.7 which shows how to obtain a Jordan canonical basis for T which also puts the scalar product in a canonical form.

The following theorem starts us off by showing that the g-spaces of a self-adjoint operator are always non-degenerate, in fact it says even more:

Lemma 3.3

Suppose $V^\mathbb{C}$ is a scalar product space and T is a real self-adjoint operator on $V^\mathbb{C}$. Let λ and μ be distinct eigenvalues of T , then $K_\lambda \perp K_\mu$, hence if $\lambda_1, \dots, \lambda_k$ is the spectrum of T then by Theorem 2.3, $V^\mathbb{C} = \bigoplus_{i=1}^k K_{\lambda_i}$.

As an immediate corollary we find that each generalized eigenspace is a non-degenerate subspace. □

PROOF Suppose $x \in K_\lambda$ and $y \in K_\mu$. Suppose $U_\lambda^p(x) = 0$, since $\mu \neq \lambda$ Theorem 2.3 says that U_λ is a bijection when restricted to K_μ , hence there exists a $z \in K_\mu$ such that $y = U_\lambda^p(z)$. Since U_λ^p is self-adjoint, property 2 implies that $\langle x, y \rangle = 0$.

Thus $K_\lambda \perp K_\mu$. As a consequence of this and Theorem 2.3 we see that $E = K_\lambda \oplus K_\lambda^\perp$, hence K_λ is non-degenerate. ■

Suppose V is a scalar product space and T is a self-adjoint operator on V . Suppose λ is an eigenvalue of T and $x \in K_\lambda$ generates a cycle of generalized eigenvectors of T of length p . Let $U = (T - \lambda I)$ and $v_i = U^{p-i}x$ for $i \in \{1, \dots, p\}$. Then observe that

$$\begin{aligned}\langle v_i, v_j \rangle &= \langle U^{p-i}x, U^{p-j}x \rangle \\ &= \langle U^{2p-i-j}x, x \rangle\end{aligned}\tag{3.1}$$

If $i + j \leq p$ then by property 2 and the fact that $U^p x = 0$ the above equation implies that $\langle v_i, v_j \rangle = 0$. If $i + j > p$ then the above equation implies that $\langle v_i, v_j \rangle$ only depends on the sum $i + j$. Thus in a cycle of length p there are only p scalar products that are variable and the above equation shows us that we only need to deal with the products $\langle v_i, v_p \rangle$. The following lemma will show that for every g-space we can always find a generator of a cycle such that $\langle v_1, v_p \rangle \neq 0$.

Lemma 3.4

Suppose V is a scalar product space and T is a self-adjoint operator on V . Fix an eigenvalue λ of T and let $U = (T - \lambda I)|_{K_\lambda}$. Suppose $k \geq 0$ satisfies $U^k \neq 0$, then there exists an $x \in K_\lambda$ such that $\langle U^k x, x \rangle \neq 0$. \square

PROOF Suppose to the contrary that $\langle U^k(x), x \rangle = 0$ for all $x \in K_\lambda$. Define a bilinear form $[\cdot, \cdot] : K_\lambda \times K_\lambda \rightarrow \mathbb{F}$ by $[x, y] = \langle U^k(x), y \rangle$. Since U^k is self-adjoint, $[\cdot, \cdot]$ is a symmetric bilinear form. Thus by the polarization identity, it follows that for any $x, y \in K_\lambda$

$$0 = [x, y] = \langle U^k(x), y \rangle$$

Now, since $U^k \neq 0$, there exists an $x \in K_\lambda$ such that $U^k x \neq 0$. But by Lemma 3.3 the scalar product is non-degenerate, hence the above equation implies that $U^k x = 0$, a contradiction. Hence the conclusion holds. \blacksquare

Assuming $\langle v_1, v_p \rangle \neq 0$, the following proposition shows how to adapt the cycle so that any other remaining scalar products are zero.

Lemma 3.5

Suppose the v_i are as defined as above for a cycle of generalized eigenvectors of T generated by $x \in K_\lambda$. Let $H \subseteq K_\lambda$ be the subspace corresponding to the cycle generated by x . If $\langle v_1, v_p \rangle \neq 0$, then we can choose an $x' \in H$ such that x' generates a cycle of generalized eigenvectors $v'_i = U^{p-i}x'$ of length p spanning H , such that v_1, \dots, v_p forms a skew-normal sequence of sign $\text{sgn} \langle v_1, v_p \rangle$ if $\lambda \in \mathbb{R}$ or 1 if $\lambda \in \mathbb{C} \setminus \mathbb{R}$. \square

PROOF Suppose first that $\lambda \in \mathbb{C} \setminus \mathbb{R}$, then let $v'_p = \langle v_1, v_p \rangle^{-\frac{1}{2}} v_p$ where any square root is fine. Then observe that:

$$\begin{aligned}\langle v'_1, v'_p \rangle &= \langle v_1, v_p \rangle^{-1} \langle v_1, v_p \rangle \\ &= 1\end{aligned}$$

If $\lambda \in \mathbb{R}$ then let $v'_p = |\langle v_1, v_p \rangle|^{-\frac{1}{2}} v_p$. Then observe that:

$$\begin{aligned}\langle v'_1, v'_p \rangle &= |\langle v_1, v_p \rangle|^{-1} \langle v_1, v_p \rangle \\ &= \pm 1\end{aligned}$$

Thus we can assume that $|\langle v_1, v_p \rangle| = 1$. Inductively suppose that $|\langle v_1, v_p \rangle| = 1$ and that $\langle v_i, v_p \rangle = 0$ for $2 \leq i \leq k-1$ for some $k \geq 2$.

Let $v'_p = v_p + av_{p-k+1}$ where a is to be determined. Now for $i \in \{1, \dots, p\}$

$$\begin{aligned}v'_i &= U^{p-i} v'_p \\ &= U^{p-i} v_p + a U^{p-i} v_{p-k+1} \\ &= v_i + av_{i-k+1}\end{aligned}$$

Observe that $v'_i = v_i$ if $i - k + 1 \leq 0$, i.e. $i \leq k - 1$. The above equation also shows that each $v'_i \in H$ and since $v'_1 = v_1 \neq 0$ the cycle generated by v'_p has length p and thus forms a basis for H . Now using the fact that $\langle v_i, v_j \rangle$ only depends on $i + j$, we find that:

$$\begin{aligned}\langle v'_k, v'_p \rangle &= \langle v_k + av_1, v_p + av_{p-k+1} \rangle \\ &= \langle v_k, v_p \rangle + a \langle v_k, v_{p-k+1} \rangle + a \langle v_1, v_p \rangle + a^2 \langle v_1, v_{p-k+1} \rangle \\ &= \langle v_k, v_p \rangle + 2a \langle v_1, v_p \rangle\end{aligned}$$

where $\langle v_1, v_{p-k+1} \rangle = 0$ since $p - k + 2 \leq p$. Thus let $a = -\frac{\langle v_k, v_p \rangle}{2 \langle v_1, v_p \rangle}$ which forces $\langle v'_k, v'_p \rangle = 0$.

Now suppose $1 \leq i < k$, then note that $v'_i = v_i$, thus

$$\begin{aligned}\langle v'_i, v'_p \rangle &= \langle v_i, v_p + av_{p-k+1} \rangle \\ &= \langle v_i, v_p \rangle + a \langle v_i, v_{p-k+1} \rangle \\ &= \langle v_i, v_p \rangle\end{aligned}$$

where $\langle v_i, v_{p-k+1} \rangle = 0$ follows from the induction hypothesis in conjunction with the fact that because $k \geq 2$, we have that $p + i - k + 1 \leq p + k - 1$ and $k \leq p$ implies $p + i - k + 1 \neq 1$. Thus v'_p satisfies the induction hypothesis and after relabeling v'_p as v_p we can apply the induction hypothesis again until $k = p$ in which case we will have proven the statement. \blacksquare

Suppose x generates a cycle of generalized eigenvectors satisfying the conclusions of the above proposition and let $z_i = U^{p-i}x$. Then by Eq. (3.1) we find that the only non-zero scalar products are $\langle z_i, z_j \rangle = \langle z_1, z_p \rangle$ where $i + j = p + 1$. Thus we say a

given cycle of generalized eigenvectors with eigenvalue λ for a self-adjoint operator are adapted to the scalar product, if they form a skew-normal sequence of sign ± 1 if $\lambda \in \mathbb{R}$ or sign 1 if $\lambda \in \mathbb{C} \setminus \mathbb{R}$. If $\lambda \in \mathbb{R}$, then $\{z_1, \dots, z_p\}$ form a real basis for the Jordan canonical form of T . If $\lambda \in \mathbb{C} \setminus \mathbb{R}$ (WLOG we can assume $\text{Im}(\lambda) > 0$), then we choose a canonical real basis $\{u_1, v_1, \dots, u_p, v_p\}$ for T as follows. Let

$$u_i = \frac{1}{\sqrt{2}}(z_i + \bar{z}_i) \quad (3.3a)$$

$$v_i = \frac{1}{i\sqrt{2}}(z_i - \bar{z}_i) \quad (3.3b)$$

Note that

$$\begin{aligned} \langle u_i, u_j \rangle &= \frac{1}{2}(\langle z_i, z_j \rangle + \langle \bar{z}_i, \bar{z}_j \rangle) \\ \langle v_i, v_j \rangle &= \frac{-1}{2}(\langle z_i, z_j \rangle + \langle \bar{z}_i, \bar{z}_j \rangle) \\ \langle u_i, v_j \rangle &= \frac{1}{2i}(\langle z_i, z_j \rangle - \langle \bar{z}_i, \bar{z}_j \rangle) = 0 \end{aligned}$$

It then follows that $\langle u_i, u_j \rangle = 1 = -\langle v_i, v_j \rangle$ if $i + j = p + 1$ with all other scalar products zero. Hence $\{u_i\}$ (resp. $\{v_i\}$) form a skew-normal sequence of sign 1 (resp. -1). Now if we set $u_{p+1} = v_{p+1} = 0$, T acts on this basis as follows:

$$\begin{aligned} Tu_i &= \frac{1}{\sqrt{2}}(\lambda z_i + z_{i+1} + \bar{\lambda} \bar{z}_i + \bar{z}_{i+1}) \\ &= \frac{1}{\sqrt{2}}((a + ib)z_i + (a - ib)\bar{z}_i) + u_{i+1} \\ &= au_i + \frac{b}{i\sqrt{2}}(\bar{z}_i - z_i) + u_{i+1} \\ &= au_i - bv_i + u_{i+1} \end{aligned}$$

Similarly

$$\begin{aligned} Tv_i &= \frac{1}{i\sqrt{2}}(\lambda z_i + z_{i+1} - \bar{\lambda} \bar{z}_i - \bar{z}_{i+1}) \\ &= \frac{1}{i\sqrt{2}}((a + ib)z_i - (a - ib)\bar{z}_i) + v_{i+1} \\ &= av_i + \frac{b}{\sqrt{2}}(z_i + \bar{z}_i) + v_{i+1} \\ &= av_i + bu_i + v_{i+1} \end{aligned}$$

In the following proposition we use these basis to show that the real subspace spanned by an adapted p -cycle is non-degenerate.

Lemma 3.6

Suppose V is a real scalar product space and T is a self-adjoint operator on V . Let x be a generator for a p -cycle of generalized eigenvectors for T with eigenvalue λ adapted to the scalar product. Let $z_i = U^{p-i}x$, H be the real subspace spanned by this cycle and $\epsilon = \langle z_1, z_p \rangle = \pm 1$. Then H is non-degenerate.

Furthermore if $\lambda \in \mathbb{R}$, then

$$\begin{aligned} \dim H &= p \\ \text{ind } H &= \begin{cases} \lfloor \frac{p+1}{2} \rfloor & \text{if } \epsilon = -1 \\ p - \lfloor \frac{p+1}{2} \rfloor & \text{if } \epsilon = 1 \end{cases} \end{aligned}$$

If $\lambda \in \mathbb{C} \setminus \mathbb{R}$, then

$$\begin{aligned} \dim H &= 2p \\ \text{ind } H &= p \end{aligned}$$

□

PROOF If $\lambda \in \mathbb{R}$, then the result follows by Lemma 2.1 applied to z_1, \dots, z_p . If $\lambda \in \mathbb{C} \setminus \mathbb{R}$, consider the real vectors $\{u_1, v_1, \dots, u_p, v_p\}$ defined in Eqs. (3.3a) and (3.3b). The result follows by Lemma 2.1 applied to the sequence u_1, \dots, u_p and then to v_1, \dots, v_p . ■

The following theorem is from [O'N83, P. 260-261].

Theorem 3.7 (Existence of the metric-Jordan canonical form [O'N83]) $\bigoplus_{i=1}^k V_i$
A linear operator T on a scalar product space V is self adjoint if and only if $V = \bigoplus_{i=1}^k V_i$ (hence each V_i is non-degenerate) where each subspace V_i is T -invariant and $T|_{V_i}$ has one of the following forms:

$$\begin{pmatrix} \lambda & 1 & & \\ & \lambda & \ddots & 0 \\ & & \ddots & 1 \\ & & & \lambda & 1 \\ 0 & & & & \lambda \end{pmatrix}$$

relative to a skew-normal sequence $\{v_1, \dots, v_p\}$ with all scalar products zero except $\langle v_i, v_j \rangle = \epsilon = \pm 1$ when $i + j = p + 1$, or

$$\begin{pmatrix} a & b & 1 & 0 & & \\ -b & a & 0 & 1 & & 0 \\ & & \ddots & & & \\ & & & a & b & 1 & 0 \\ & & & -b & a & 0 & 1 \\ 0 & & & & & a & b \\ & & & & & -b & a \end{pmatrix}$$

relative to a basis $\{u_1, v_1, \dots, u_p, v_p\}$ with all scalar products zero except $\langle u_i, u_j \rangle = 1 = -\langle v_i, v_j \rangle$ if $i + j = p + 1$.

The index and dimension of V_i is determined by the blocks $T|_{V_i}$ due to Lemma 3.6, hence we must have $\text{ind } V = \sum_{i=1}^k \text{ind } V_i$ and $n = \sum_{i=1}^k \dim V_i$. \square

PROOF We proceed by induction. If $n = 1$ then this result trivially holds. So suppose $n \geq 2$ and this result is true for all self-adjoint operators on scalar product spaces of dimension strictly less than n . Now we show that this holds when $\dim V = n$.

Fix an eigenvalue λ for T (which exists after complexification of V if necessary). Let $U = (T - \lambda I)$. Let p be the smallest integer such that $\dim N(U^p) = \dim N(U^{p+1})$, thus $K_\lambda = N(U^p)$. Then $\dim N(U^{p-1}) < \dim N(U^p)$, hence $U^{p-1}|_{K_\lambda} \neq 0$, thus by Lemma 3.4 there exists an $x \in K_\lambda$ such that $\langle U^{p-1}x, x \rangle \neq 0$. Note that by construction p is the smallest integer such that $U^p x = 0$, hence x generates a p -cycle of generalized eigenvectors with eigenvalue λ .

Hence by Lemma 3.5, the p -cycle of generalized eigenvectors generated by x can be modified into another such p -cycle spanning the same subspace as the original and adapted to the scalar product. Thus we now assume that the p -cycle of generalized eigenvectors generated by x is adapted to the scalar product. Note that it follows by Lemma 2.1 that the set of p vectors in this cycle are linearly independent. Let H be the real subspace spanned by the p -cycle(s) generated by x if $\lambda \in \mathbb{R}$ or by x and its conjugate if $\lambda \in \mathbb{C} \setminus \mathbb{R}$. By Lemma 3.6, H is non-degenerate and by construction H is T -invariant. If $H = V$ then we are done, so assume $H \subsetneq V$. Then by property 1, H^\perp is an invariant subspace of T , and is complementary to H by non-degeneracy of H . Let $T' = T|_{H^\perp}$, then H^\perp is a scalar product space with $0 < \dim H^\perp < n$ and T' is a self-adjoint operator on H^\perp . Hence the induction hypothesis applies to T' , in which case we conclude that the result holds for T .

The converse is also easily checked. \blacksquare

4 Uniqueness of the metric-Jordan canonical form

In this section T is self adjoint operator on a scalar product space V . We will show in what sense each self-adjoint operator T admits a “unique” metric-Jordan canonical form. We will do this by showing that the parameters appearing in any two canonical forms derived by Theorem 3.7 must be the same. Then we will show how this result can be used to determine if two self-adjoint operators are isometrically equivalent.

Lemma 4.1

Let $U = (T - \lambda I)$ for some eigenvalue λ , suppose x generates an adapted cycle of length p and sign ε and denote by $v_i = U^{p-i}x$. Also let H be the subspace spanned by this cycle.

For any $0 \leq k \leq p-1$ define a symmetric bilinear form $[\cdot, \cdot]_k$ on H by

$$[x, y]_k = \langle U^k x, y \rangle$$

for $x, y \in H$. Then the number of zeros in any diagonal representation for $[\cdot, \cdot]_k$ is k . If the $\lambda \in \mathbb{R}$ then the number of negative entries in any diagonal representation for $[\cdot, \cdot]_k$ is

$$\begin{cases} \lfloor \frac{(p-k)+1}{2} \rfloor & \text{if } \epsilon = -1 \\ (p-k) - \lfloor \frac{(p-k)+1}{2} \rfloor & \text{if } \epsilon = 1 \end{cases}$$

In conclusion, we see that the of invariants of $[\cdot, \cdot]_k$ depends only on p, k, ϵ . \square

PROOF We prove this by exhibiting a diagonal representation for $[\cdot, \cdot]$ restricted to H .

First observe that for $i, j \in \{1, \dots, p\}$

$$[v_i, v_j]_k = \langle U^k U^{p-i} x, U^{p-j} x \rangle = \langle U^{2p+k-i-j} x, x \rangle$$

The above equation is non-zero iff

$$\begin{aligned} 2p + k - i - j &= p - 1 \\ \Leftrightarrow p + k - i - j &= -1 \\ \Leftrightarrow i + j &= p + k + 1 \end{aligned}$$

It follows that if $i < k + 1$, then $[v_i, v_j] = 0$ for any j . Now define vectors $v'_i = v_{i+k}$ for $i \in \{1, \dots, p - k\}$. Then $\langle v'_i, v'_j \rangle \neq 0$ iff

$$\begin{aligned} i + k + j + k &= p + k + 1 \\ \Leftrightarrow i + j &= p - k + 1 \end{aligned}$$

Hence v'_1, \dots, v'_{p-k} (or equivalently v_{k+1}, \dots, v_p) form a pseudo-orthonormal set of vectors with sign ϵ . Thus the formula for the number of negative entries when $\lambda \in \mathbb{R}$ follows from Lemma 2.1. Also observe that the number of zeros is k . Then by Sylvester's law of inertia it follows that the invariants of $[\cdot, \cdot]_k$ are given as above and hence depend only on p, k, ϵ . \blacksquare

For a real eigenvalue λ , an adapted cycle $x, Ux, \dots, U^{p-1}x$ is called *positive* if $\langle U^{p-1}x, x \rangle = 1$ or negative if $\langle U^{p-1}x, x \rangle = -1$. By a metric-Jordan canonical basis, we mean one that is obtained from Theorem 3.7.

Theorem 4.2 (Uniqueness of the metric-Jordan canonical form)

Suppose λ is an eigenvalue of T . If $\lambda \in \mathbb{R}$, then the number of positive (negative) cycles in K_λ of a given length is independent of any metric-Jordan canonical basis. If $\lambda \in \mathbb{C} \setminus \mathbb{R}$, then the number of cycles in K_λ of a given length is independent of any metric-Jordan canonical basis. \square

PROOF Fix an eigenvalue λ of T and let $U = (T - \lambda I)$. Restrict the argument to the vector space K_λ , i.e set $V = K_\lambda$. Denote by $[\cdot, \cdot]_i$ the symmetric bilinear form given by:

$$[x, y]_i = \langle U^i x, y \rangle$$

for $x, y \in K_\lambda$. We will prove that the number of positive (negative) cycles of a given length depend only on the number of positive (negative) entries in a diagonal representation for $[\cdot, \cdot]_0, [\cdot, \cdot]_1, \dots, [\cdot, \cdot]_n$. It is understood that the complex representations are chosen so that there are only positive or zero entries in it. It will follow by Sylvester's law of inertia that these signs are independent of any basis.

Fix a metric-Jordan canonical basis for $T|_{K_\lambda}$. It's known that $U^l = 0$ for any $l > n$, hence it follows that the number of cycles of length larger than n are determined by invariants of $[\cdot, \cdot]_l$ for $l > n$. Suppose inductively that the statement holds for all cycles of length strictly larger than p . We will now prove the statement for cycles of length p .

Denote by H the T -invariant non-degenerate (possibly zero) subspace spanned by all cycles of length strictly larger than p in this canonical basis. Observe that since H is T -invariant, it follows for any $l \geq 0$ that $[x, y]_l = 0$ for $x \in H$ and $y \in H^\perp$.

Case 1 There are no cycles of length p in this canonical basis.

Then note that $[x, y]_{p-1} \equiv 0$ for any $x, y \in H^\perp$ and if $H \neq 0$ the invariants of $[\cdot, \cdot]_{p-1}$ on H are uniquely determined by invariants of $[\cdot, \cdot]_l$ for $l \geq p$ by Lemma 4.1. Also the invariants of $[\cdot, \cdot]_{p-1}$ over K_λ are determined by Sylvester's law of inertia, hence it follows that the number of cycles of length p are uniquely determined.

Case 2 Let x_1, \dots, x_m be generators for cycles of length p in this canonical basis. For vectors from H^\perp in this canonical basis the only non-zero diagonal entries of $[\cdot, \cdot]_{p-1}$ are

$$[x_i, x_i]_{p-1} = \langle U^{p-1} x_i, x_i \rangle = \pm 1 \quad i = 1, \dots, m$$

Again, if $H \neq 0$ the invariants of $[\cdot, \cdot]_{p-1}$ on H are uniquely determined by invariants of $[\cdot, \cdot]_l$ for $l \geq p$ by Lemma 4.1. The invariants of $[\cdot, \cdot]_{p-1}$ over K_λ are determined by Sylvester's law of inertia, hence it follows that the number of positive (and negative) cycles of length p are uniquely determined.

Thus the result follows by induction on p . ■

We can now state what we mean by “the” metric-Jordan canonical form:

Definition 4.3

Let T be a self-adjoint operator on a scalar product space V . To each adapted p -cycle of sign ε with eigenvalue $\lambda \in \mathbb{C}$ we associate a 3-tuple $(\lambda, p, \varepsilon)$. A canonical form given

by Theorem 3.7 gives an un-ordered list of such 3-tuples counting multiplicities. We call this list *the metric-Jordan canonical form*. \square

By the above theorem, it follows that the above definition is well defined, i.e. each self-adjoint operator T admits precisely one metric-Jordan canonical form. The following example shows that the signs appearing in these canonical forms add some subtleties:

Example 4.4

Suppose V is Minkowski space equipped with the standard metric

$$g = \text{diag}(-1, 1, \dots, 1)$$

For $\lambda_1 < \dots < \lambda_n \in \mathbb{R}$ define two self-adjoint operators T_1 and T_2 as follows:

$$T_1 = \text{diag}(\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_n)$$

$$T_2 = \text{diag}(\lambda_2, \lambda_1, \lambda_3, \dots, \lambda_n)$$

Now observe that even though T_1 and T_2 have the same eigenvalues, they have different metric-Jordan canonical forms. We will show shortly that T_1 and T_2 are isometrically inequivalent, in the sense that there is no $R \in O(V)$ which relates T_1 and T_2 by a similarity transformation. \square

Note that the above example is in sharp contrast with the Euclidean case where T_1 and T_2 as defined above would be isometrically equivalent.

Theorem 4.5 (Isometric Equivalence of self-adjoint operators)

Suppose S and T are self-adjoint operators on a scalar product space V . Then S and T differ by an isometry $R \in O(V)$ iff they have the same metric-Jordan canonical form. \square

PROOF It's clear that if S and T have the same metric-Jordan canonical form then there is an isometry $R \in O(V)$ which relates the two operators, namely the transformation that relates a metric-Jordan canonical basis of S to a metric-Jordan canonical basis of T .

Suppose T is given as follows relative to S :

$$T = RSR^{-1}$$

Let $\beta = \{v_1, \dots, v_n\}$ be a canonical basis for S . Then consider the basis $\tilde{\beta} = \{Rv_1, \dots, Rv_n\}$ for T . Since R is an isometry, we have

$$g|_{\tilde{\beta}} = g|_{\beta}$$

The equation relating T to S implies that

$$T|_{\tilde{\beta}} = S|_{\beta}$$

Hence S and T have the same metric-Jordan canonical form. \blacksquare

5 Minkowski Space

Fix a self-adjoint operator T in Minkowski space. We will use Theorem 3.7 to enumerate the possible Jordan canonical forms of T together with the metric in an adapted basis. As a consequence of Theorem 3.7, we simply have to determine which combination of Jordan blocks are possible in Minkowski space by imposing the dimension and signature restrictions. This can be done with the help of Lemma 3.6, since it gives us the index of a given subspace associated with a Jordan block. We denote by D_k a diagonal $k \times k$ matrix and I_k the identity $k \times k$ matrix. We have the following cases.

Case 1 T is diagonalizable with real spectrum

In this case T must have a time-like eigenvector. Indeed, since each eigenspace E_λ is non-degenerate, one eigenspace, say H , must have index 1. Then by obtaining an orthonormal basis for H , we can obtain a time-like eigenvector. Thus T has the following form:

$$T = D_n \qquad g = \text{diag}(-1, 1, \dots, 1)$$

Case 2 T has a complex eigenvalue $\lambda = a + ib$ with $b \neq 0$

By Lemma 3.6 the real subspace H spanned by a complex eigenvector with eigenvalue λ and its complex conjugate must have index 1. Since this subspace is T -invariant, by Item 1 H^\perp is a complementary invariant subspace, which must be Euclidean. Hence T must have the following form:

$$T = \begin{pmatrix} a & b & 0 \\ -b & a & \\ 0 & & D_{n-2} \end{pmatrix} \qquad g = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & \\ 0 & & I_{n-2} \end{pmatrix}$$

Case 3 T has real eigenvalues but is not diagonalizable

In this case we go through the possible multidimensional Jordan blocks associated to a real irreducible subspace, say H , of T . By Theorem 3.7, each basis for this subspace can be adapted to the scalar product, hence is non-degenerate. By Lemma 3.6 there are three types of Jordan blocks which have an associated subspace, H , with index one. For each of these subspaces, H^\perp is a complementary T -invariant Euclidean subspace. The first two cases occur when $\dim H = 2$, and are given as follows:

$$T = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & \\ 0 & & D_{n-2} \end{pmatrix} \qquad g = \begin{pmatrix} 0 & \epsilon & 0 \\ \epsilon & 0 & \\ 0 & & I_{n-2} \end{pmatrix} \quad \epsilon = \pm 1$$

Note that the above form contains two metric-Jordan canonical forms depending on the sign of ϵ . The third occurs when $\dim H = 3$:

$$T = \begin{pmatrix} \lambda & 1 & 0 & & \\ 0 & \lambda & 1 & & 0 \\ 0 & 0 & \lambda & & \\ & & & D_{n-3} & \\ 0 & & & & \end{pmatrix} \quad g = \begin{pmatrix} 0 & 0 & 1 & & \\ 0 & 1 & 0 & & 0 \\ 1 & 0 & 0 & & \\ & & & I_{n-3} & \\ 0 & & & & \end{pmatrix}$$

We also note that this case (T has real eigenvalues but is not diagonalizable) holds iff T has a unique lightlike eigenvector. This fact can be deduced by inspection of the above canonical forms.

Now, we collect some necessary and sufficient conditions concerning the diagonalizability of T in the following theorem. The second and third facts are from Theorem 4.1 in [Hal+96] while the last fact is from Section 9.5 in [Gre75]. All these facts can be readily deduced from the canonical forms listed above.

Theorem 5.1 (Properties of self-adjoint operators in Minkowski space)

Let V be a Minkowski space and T a self-adjoint operator on V . Then the following statements are true:

1. *T is diagonalizable with a real spectrum iff T has 1 timelike eigenvector or equivalently T has $n - 1$ linearly independent spacelike eigenvectors.*
2. *If T has two linearly independent null eigenvectors then T is diagonalizable with a real spectrum and T has a time-like eigenspace of dimension at least 2 containing these eigenvectors.*
3. *If T has a real spectrum, then T is diagonalizable iff it has no null eigenvectors or at least two linearly independent null eigenvectors. In other words, T is not diagonalizable iff it has a unique null eigendirection.*
4. *If $n \geq 3$ and $\langle Tx, x \rangle \neq 0$ for all null vectors x then T is diagonalizable with a real spectrum.* □

References

- [Bro06] T.J.I.A. Bromwich. *Quadratic Forms and Their Classification by Means of Invariant-factors*. Cambridge tracts in mathematics and mathematical physics. University Press, 1906 (cit. on p. [1](#)).
- [FIS03] S H Friedberg, A J Insel, and L E Spence. *Linear Algebra*. Pearson Education, 2003. ISBN: 9780131202665 (cit. on pp. [5](#), [6](#)).
- [GL05] Israel Gohberg and P. Lancaster. *Indefinite linear algebra and applications*. Basel: Birkhauser, 2005, p. 357. ISBN: 9783764373498 (cit. on p. [1](#)).
- [Gre75] W H Greub. *Linear Algebra*. 3rd ed. Springer, 1975, p. 476. ISBN: 0387901108 (cit. on p. [17](#)).
- [Hal+96] G S Hall et al. “On the algebraic structure of second order symmetric tensors in 5-dimensional space-times”. In: *General Relativity and Gravitation* 28.9 (1996), pp. 1107–1113. ISSN: 0001-7701. DOI: [10.1007/BF02113160](#) (cit. on p. [17](#)).
- [Jac12] N Jacobson. *Basic Algebra I*. Dover books on mathematics. Dover Publications, Incorporated, 2012. ISBN: 9780486135229 (cit. on p. [6](#)).
- [LR05] P. Lancaster and L. Rodman. “Canonical forms for symmetric/skew-symmetric real matrix pairs under strict equivalence and congruence”. In: *Linear Algebra and its Applications* 406 (Sept. 2005), pp. 1–76. ISSN: 00243795. DOI: [10.1016/j.laa.2005.03.035](#) (cit. on p. [1](#)).
- [O’N83] B. O’Neil. *Semi-Riemannian Geometry: With Applications to Relativity*. 3rd ed. Pure and Applied Mathematics Series v. 103. Acad. Press, 1983, p. 468. ISBN: 9780125267403 (cit. on pp. [ii](#), [1](#), [3](#), [7](#), [11](#)).